# Structure preserving Stochastic Impulse Methods for stiff Langevin systems with a uniform global error of order 1 or 1/2 on position

Molei Tao[†§], Houman Owhadi[†‡], Jerrold E. Marsden[†‡]

June 25, 2010

### Abstract

Impulse methods are generalized to a family of integrators for Langevin systems with quadratic stiff potentials and arbitrary soft potentials. Uniform error bounds (independent from stiff parameters) are obtained on integrated positions allowing for coarse integration steps. The resulting integrators are explicit and structure preserving (quasi-symplectic for Langevin systems).

## 1 Introduction

**Results:** This paper generalizes the impulse methods for stiff Hamiltonian systems [15, 39] to stiff stochastic Langevin systems. In the stochastic setting these integrators are quasi-symplectic as defined in [29]. Uniform error bounds are obtained for both stochastic and deterministic settings.

More precisely, this paper is concerned with the numerical integration the following stiff SDEs:

$$\begin{cases} M\,dq &=& p\,dt \\ dp &=& -\nabla V(q)dt - \epsilon^{-1}Kq\,dt - cp\,dt + \sigma\,dW \end{cases}$$

which describes a stochastic mechanical system with a potential being sum of slow $V(q)$ and fast $\frac{1}{2}\epsilon^{-1}q^T Kq$, and the momentum being perturbed by noise and attenuated by friction.

When noise and friction are both present (i.e. $c \neq 0$, $\sigma \neq 0$), a 1st-order member of the proposed Stochastic Impulse Methods (SIMs) family will integrate position $q$ with a global error uniformly bounded by $CH^{1/2}$, where $H$ is the integration timestep and $C$ is a constant independent from $\epsilon^{-1}$ (provided that the solution remains bounded). The integrator is also shown to be quasi-symplectic. When noise and friction are absent (i.e. $c = 0$, $\sigma = 0$), the deterministic (and symplectic) version of the 1st-order SIM gives a uniform 1st-order global error on $q$, if again the solution is bounded. The error bound on momentum $p$, however, is not uniform here. Recall that increased accuracy and stability was one of motivations supporting the development of mollified impulse methods [12, 32].

*Dynamical systems with multiple time scales pose a major problem in simulations because the small time steps required for stable integration of the fast motions lead to large numbers of time*

---

[†]Control & Dynamical Systems, MC 107-81,

[‡]Applied & Computational Mathematics, MC 217-50, California Institute of Technology, Pasadena, CA 91125, USA

[§]Corresponding author; Email: mtao@caltech.edu

*steps required for the observation of slow degrees of freedom* [16]. As seen from the error bounds, in the case of quadratic fast potential, SIMs provide a possibility of accurate integration with a choice of timestep not restricted by the stiffness $\epsilon^{-1}$, as long as position is the quantity of interest. In these cases, a large timestep can be adopted.

Also, SIMs are symplectic [16] and in fact variational [26] in the case of no noise no friction, and are quasi-symplectic [29] in the case of full Langevin. As a result of the preservation of structure, properties such as near preservation of energy or of associated Boltzmann-Gibbs invariant measure, as well as conservation of momentum maps could be obtained, which significantly benefit long time numerical integrations.

**Related work:** Many elegant methods have been proposed in the area of stiff Hamiltonian/Langevin integration with different focus and perspective.

Impulse methods, as well as other members of the exponential integrator family [14], including Mollified Impulse Methods [12, 32] and Gautschi-type integrators [17] are prevailing symplectic integrators for stiff Hamiltonian systems. They are however not directly extendable to stiff Langevin systems if integration with a large step is desired. The general GLA [3] approach (see also [34] for an extension of impulse methods non-stiff Langevin systems) of constructing Langevin integrator from a symplectic scheme by composing an Ornstein-Uhlenbeck flow with the symplectic integrator will not yield a uniform error bound in the case of stiff Langevins. It is worth mentioning that impulse methods are not limited to quadratic stiff potentials (provided that the flow of the stiff part of the Hamiltonian is given).

The implicit method approach for integrating stiff equations include the LIN algorithm [40] for stiff Langevin systems. However, it has been observed that *Implicit methods in general fail to capture the effective dynamics of the slow time scale because they cannot correctly capture non-Dirac invariant distributions* [24]. Moreover implicit methods are generally slower than explicit methods, provided they use comparable timesteps.

Implicit and explicit approaches were combined in a variational integration framework by defining the discrete Lagrangian via trapezoidal approximation of the soft potential and midpoint approximation of the stiff potential. The resulting IMEX for stiff Hamiltonian systems [35] is explicit in the case of quadratic fast potential. Similar as the case of impulse methods, there is no easy way to extend IMEX to stiff Langevin systems.

The Hamilton-Jacobi derived homogenization method for multiscale Hamiltonian systems [22] enables the usage of a large timestep for deterministic systems but can not directly be extended to stiff Langevin systems using the GLA approach [3].

Multiscale methods that integrate the slow dynamics by averaging the effective contribution of the fast dynamics have been applied to stiff Langevin systems. These include Heterogeneous Multiscale Methods (HMM) [8, 10, 7, 4, 1], equation free methods [19, 13, 20], and FLow AVeraging integratORS (FLAVORS) [37]. We observe that these methods use mesoscopic timesteps, which are (usually) one or two orders of magnitude smaller than the large steps employed in SIMs. These methods also assume a separation of timescales, and therefore will not work for generic stiff Langevin systems that are not necessarily multiscale. In addition, based on averaging instantaneous drifts, both Heterogeneous Multiscale Methods and equation free methods (in their original form) require an identification of slow variables in general nonlinear cases (with exceptions such as in [9]). Reversible and symmetric methods in these frameworks have been proposed [23, 2, 33]. FLAVORS are based on averaging instantaneous flows and do not require explicit identification of slow variables, and are symplectic (quasi-symplectic).

# 2 Stochastic Impulse Methods

Consider numerical integration of the following multiscale Langevin SDEs

$$\begin{cases} Mdq &= pdt \\ dp &= -\nabla V(q)dt - \epsilon^{-1}Kqdt - cpdt + \sigma dW \end{cases} \tag{1}$$

where $0 < \epsilon \ll 1$, $q \in \mathbb{R}^d$, $p \in \mathbb{R}^d$, $K$ is positive definite $d \times d$ matrix, $c$ and $\sigma$ are positive semi-definite $d \times d$ matrices, respectively indicating viscous damping coefficients and amplitudes of noises. We restrict ourselves to Euclidean phase spaces, although the method is readily generalizable to manifolds. In addition, we require that matrices $K$ and $c$ commute; a special case satisfying this requirement is $c$ being a scalar.

In the case of no noise no friction ($c = 0$ and $\sigma = 0$), the system degenerates to a deterministic mechanical system with Hamiltonian $H(q, p) = \frac{1}{2}p^T M^{-1}p + V(q) + \epsilon^{-1}\frac{1}{2}q^T Kq$.

Also, the method as well as the uniform convergence theorem works for a more general open system:

$$\begin{cases} Mdq &= pdt \\ dp &= F(q)dt - \epsilon^{-1}Kqdt - cpdt + \sigma dW \end{cases} \tag{2}$$

but we stick to (2) for simplicity in descriptions.

Denote by $\phi^f(\tau) : \big(q^f(t), p^f(t)\big) \mapsto \big(q^f(t + \tau), p^f(t + \tau)\big)$ and $\phi^s(\tau) : (q^s(t), p^s(t)) \mapsto (q^s(t + \tau), p^s(t + \tau))$ respectively the $\tau$-flow maps of the autonomous SDE systems

$$\begin{cases} Mdq^f &= p^f dt \\ dp^f &= -\epsilon^{-1}Kq^f dt - cp^f dt + \sigma dW \end{cases} \tag{3}$$

and

$$\begin{cases} Mdq^s &= 0 \\ dp^s &= -\nabla V(q^s)dt \end{cases} \tag{4}$$

Since the first system is a linear SDE and the second is a free drift, flows of both can be obtained exactly.

Then Stochastic Impulse Methods(SIMs) are defined via compositions of $\phi^f$ and $\phi^s$. Here are several examples of SIMs with a timestep $H$:

**Integrator 1.** *1st order SIM in the* $c = 0$, $\sigma = 0$ *case, is given by the one step update of* $\phi^s(H) \circ \phi^f(H)$:

$$\begin{cases} q_{k'} &= A_{11}(H)q_k + A_{12}(H)p_k \\ p_{k'} &= A_{21}(H)q_k + A_{22}(H)p_k \\ q_{k+1} &= q_{k'} \\ p_{k+1} &= p_{k'} - H\nabla V(q_{k'}) \end{cases}$$

$$where \; \begin{bmatrix} A_{11}(H) & A_{12}(H) \\ A_{21}(H) & A_{22}(H) \end{bmatrix} = \exp \begin{bmatrix} 0 & M^{-1}H \\ -\epsilon^{-1}KH & 0 \end{bmatrix},$$

$$\begin{cases} q_0 = q(0) \\ p_0 = p(0) \end{cases}$$

**Remark 2.1.** The other 1st order SIM, as the above's dual, can be obtained via the one step update $\phi^f(H) \circ \phi^s(H)$. Both these 1st order composition schemes are well known as the Lie-Trotter splitting [38].

**Integrator 2.** *1st order SIM in the full Langevin case*, *given by the same one step update* $\phi^s(H) \circ \phi^f(H)$:

$$\begin{cases} \begin{array}{rcl} q_{k'} & = & B_{11}(H)q_k + B_{12}(H)p_k + Rq_k(H) \\ p_{k'} & = & B_{21}(H)q_k + B_{22}(H)p_k + Rp_k(H) \\ q_{k+1} & = & q_{k'} \\ p_{k+1} & = & p_{k'} - H\nabla V(q_{k'}) \\ \begin{bmatrix} Rq_k(H) \\ Rp_k(H) \end{bmatrix} & \sim & \mathcal{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{11}^2(H)) & \Sigma_{12}^2(H)) \\ \Sigma_{21}^2(H)) & \Sigma_{22}^2(H)) \end{bmatrix}), i.i.d. \; normal \; distributed \end{array} \end{cases}$$

$$where \; \begin{bmatrix} B_{11}(H) & B_{12}(H) \\ B_{21}(H) & B_{22}(H) \end{bmatrix} = \exp \begin{bmatrix} 0 & M^{-1}H \\ -\epsilon^{-1}KH & -cH \end{bmatrix},$$

$$\begin{cases} q_0 = q(0) \\ p_0 = p(0) \end{cases},$$

$$\begin{cases} \Sigma_{11}^2(H) = \int_{s=0}^H \left( B_{12}(H-s)\sigma\sigma^T B_{12}^T(H-s) \right) ds \\ \Sigma_{12}^2(H) = \int_{s=0}^H \left( B_{12}(H-s)\sigma\sigma^T B_{22}^T(H-s) \right) ds \\ \Sigma_{21}^2(H) = \int_{s=0}^H \left( B_{22}(H-s)\sigma\sigma^T B_{12}^T(H-s) \right) ds \\ \Sigma_{22}^2(H) = \int_{s=0}^H \left( B_{22}(H-s)\sigma\sigma^T B_{22}^T(H-s) \right) ds \end{cases}$$

**Remark 2.2.** $\begin{bmatrix} Rq_k(H) \\ Rp_k(H) \end{bmatrix}$ indicates the value of $\int_{s=0}^H B(H-s) \begin{bmatrix} 0 \\ \sigma dW_s \end{bmatrix}$ and hence is a vectorial normal random variable with zero mean and covariance of $\begin{bmatrix} \Sigma_{11}^2(H) & \Sigma_{12}^2(H) \\ \Sigma_{21}^2(H) & \Sigma_{22}^2(H) \end{bmatrix}$.

**Integrator 3.** *2nd order SIM in the full Langevin case*, *given by the one step update* $\phi^s(H/2) \circ \phi^f(H) \circ \phi^s(H/2)$:

$$\begin{cases} \begin{array}{rcl} q_{k'} & = & q_k \\ p_{k'} & = & p_k - \frac{H}{2}\nabla V(q_k) \\ q_{k''} & = & B_{11}(H)q_{k'} + B_{12}(H)p_{k'} + Rq_k(H) \\ p_{k''} & = & B_{21}(H)q_{k'} + B_{22}(H)p_{k'} + Rp_k(H) \\ q_{k+1} & = & q_{k''} \\ p_{k+1} & = & p_{k''} - \frac{H}{2}\nabla V(q_{k''}) \end{array} \end{cases}$$

**Remark 2.3.** This uses the 2nd order composition scheme known as the Strang or Marchuk splitting [36, 25]. When no noise or friction, i.e. $c = 0$ and $\Sigma = 0$, the resulting integrator degenerates to the prevailing Verlet-I/r-RESPA impulse method [15, 39].

**Remark 2.4.** Higher order SIMs can be obtained systematically since generic way for constructing higher order splitting/composition schemes exists [16]. For instance a 4th order SIM is given by $\phi^s(cH/2) \circ \phi^f(cH) \circ \phi^s((1-c)H/2) \circ \phi^f((1-2c)H) \circ \phi^s((1-c)H/2) \circ \phi^f(cH) \circ \phi^s(cH/2)$ where $c = \frac{1}{2-2^{1/3}}$ [30].

# 3 Properties

## 3.1 Symplecticity

In the case of $c = 0$ and $\sigma = 0$, since $\phi^s$ and $\phi^f$ are the exact flows of Hamiltonian systems, they are symplectic. Therefore SIMs, as compositions of the two, are symplectic.

In fact, SIMs here are not only symplectic but variational, in the sense that their equations of motion are obtained as critical point of a globally defined action, which is the integral of a discrete Lagrangian. Since SIMs are based on splitting schemes, and the original system is split to two Hamiltonian systems, backward error analysis can be done via Poisson brackets [16], resulting in a global non-degenerate Hamiltonian that is exactly preserved. Then Legendre transformation gives the discrete Lagrangian and hence the variational structure.

When noise and friction are present, SIMs are quasi-symplectic for RL1 and RL2 in [29] can be easily checked to be true, i.e. they degenerate to symplectic methods if friction is set equal to zero and the Jacobian of the flow map is independent of $(q, p)$.

If in addition $c$ is isotropic, then SIMs are conformally symplectic, i.e. they preserve the precise symplectic area change associated to the flow of inertial Langevin processes [28].

## 3.2   Uniform Convergence

In the case of $c = 0$ and $\sigma = 0$, convergence of SIMs is guaranteed by the general construction of splitting schemes. In the full Langevin setting, analogous convergence results for the same splitting schemes can be easily obtained using generators of SDEs. By this approach, however, the error bound will contain the scaling factor $\epsilon^{-1}$ and therefore restrain the timestep from being large. We instead seek for uniform convergence results, i.e. error bounds that don't depend on $\omega$. It turns out such a uniform bound holds only for the position $q$ but not the momentum $p$.

**Condition 3.1.** *We will prove a uniform bound on the scaled energy norm of the global error of Integrator 2 if the following conditions hold:*

1. *Matrices $c$ and $K$ commute. A special case could be $c$ being a scalar.*

2. *$\lim_{\epsilon \to 0} \sqrt{\epsilon}\|c\|_2 \leq C$ for some constant $C$ independent of $\epsilon$, i.e. $c \leq O(\epsilon^{-1/2})$.*

3. *$\sigma$ is independent of $\epsilon^{-1}$, in the sense that $\lim_{\epsilon \to 0} \epsilon^p \|\sigma\|_2 = 0$ for any $p > 0$.*

4. *In the integration domain of interest $\nabla V(\cdot)$ is bounded and Lipschitz continuous with coefficient $L$, i.e. $\|\nabla V(a) - \nabla V(b)\|_2 \leq L\|a - b\|_2$.*

5. *Denote by $x(T) = (q(T), p(T))$ the exact solution to (2), and $x_T = (q_T, p_T)$ the discrete numerical trajectory given by Integrator 2, then $\mathbb{E}\|x(T)\|_2^2 \leq C$ and $\mathbb{E}\|x_T\|_2^2 \leq C$ for some constant $C$ independent of $\epsilon^{-1}$ but dependent on initial condition $\mathbb{E}\|\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\|_2^2$, amplitude of noise $\sigma$ and friction $c$.*

   *Note that this condition usually holds due to preservation of Boltzmann-Gibbs invariant measure, whose parameter of temperature doesn't depend on $\epsilon^{-1}$ since noise is weak, and whose energy function is usually dominated by the positive definite fast potential (implying closed energy level sets).*

**Theorem 3.1.** *If Condition 3.1 holds, the $1^{st}$ order SIM (Integrator 2) for multiscale Langevin system (2) ($c \neq 0$, $\sigma \neq 0$) has in mean square sense a uniform global error of $O(H^{1/2})$ in $q$ and a non-uniform global error of $\epsilon^{-1/2}O(H^{1/2})$ in $p$, given a fixed total simulation time $T = NH$:*

$$(\mathbb{E}\|q(T) - q_T\|_2^2)^{1/2} \leq CH^{1/2} \tag{5}$$
$$(\mathbb{E}\|p(T) - p_T\|_2^2)^{1/2} \leq \epsilon^{-1/2}\|\sqrt{K}\|_2 CH^{1/2} \tag{6}$$

*where $q(T), p(T)$ is the exact solution and $q_T, p_T$ is the numerical solution; $C$ is a positive constant independent of $\epsilon^{-1}$ but dependent on simulation time $T$, scaleless elasticity matrix $K$,*

*scaled damping coefficient $\sqrt{\epsilon}c$ (O(1)), amplitude of noise $\sigma$, slow potential energy $V(\cdot)$, and initial condition $\mathbb{E}\|\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\|_2^2$.*

*Proof.* We refer to the appendix for the proof. □

**Remark 3.1.** By looking at the proof, one can be assured that all convergence results of SIMs apply to situations where the deterministic system is in a more general form of $M\frac{d^2}{dt^2}q = -\epsilon^{-1}Kq + F(q)$, where $F(q)$ doesn't have to be $-\nabla V(q)$.

In the special case of Hamiltonian system, the same integrator gains $1/2$ more order of accuracies.

**Condition 3.2.** *We will prove a uniform bound on the scaled energy norm of the global error of Integrator 1 if the following conditions hold:*

1. *In the integration domain of interest $\nabla V(\cdot)$ is bounded and Lipschitz continuous with coefficient $L$, i.e. $\|\nabla V(a) - \nabla V(b)\|_2 \le L\|a - b\|_2$.*

2. *Denote by $x(T) = (q(T), p(T))$ the exact solution to (2) with $c = 0$ and $\sigma = 0$, and $x_T = (q_T, p_T)$ the discrete numerical trajectory given by Integrator 1, then $\|x(T)\|_2^2 \le C$ and $\|x_T\|_2^2 \le C$ for some constant $C$ independent of $\epsilon^{-1}$ but dependent on initial condition $\|\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\|_2^2$.*

   *Note that this condition usually holds due to preservation of energy, which is usually dominated by the positive definite fast potential (implying closed energy level sets).*

**Theorem 3.2.** *If Condition 3.2 holds, the $1^{st}$ order SIM (Integrator 1) for multiscale Hamiltonian system ((2) with $c = 0$, $\sigma = 0$) has a uniform global error of $O(H)$ in $q$ and a non-uniform global error of $\epsilon^{-1/2}O(H)$ in $p$, given a fixed total simulation time $T = NH$:*

$$\|q(T) - q_T\|_2 \le CH \tag{7}$$
$$\|p(T) - p_T\|_2 \le \epsilon^{-1/2}\|\sqrt{K}\|_2 CH \tag{8}$$

*where $q(T), p(T)$ is the exact solution and $q_T, p_T$ is the numerical solution; $C$ is a positive constant independent of $\epsilon^{-1}$ but dependent on simulation time $T$, scaleless elasticity matrix $K$, slow potential energy $V(\cdot)$ and initial condition $\|\begin{bmatrix} q_0 \\ p_0 \end{bmatrix}\|_2$.*

*Proof.* It follows by simplifying the proof of Theorem 3.1. □

## 3.3 Stability

As one sees from Condition 3.1 and 3.2 (as another nonlinear demonstration of Lax equivalence theorem [21]), stability is necessary for global convergence. Instability could either come from the problem itself (not all SDEs have bounded solutions in the mean square sense), or from imperfection in numerical integration schemes. Here consider the latter possibility only. It is shown that impulse methods are not unconditionally stable [12], and its improvement, mollified impulse methods, are still susceptible to instability intervals (although narrower) in a linear example [5]. Nevertheless, instability intervals of impulse method are already narrow regions; for instance, the first instability interval in the stiff example considered by [5] is $0.544 < H < 0.553$. It is intuitive that instability intervals for the stochastic case with damping or higher order schemes will not be wider. Therefore one could still choose a large timestep $H$ in SIMs without hitting the instability, by at most a few integration tryouts with slightly varied $H$'s.

# 4 Numerical Examples

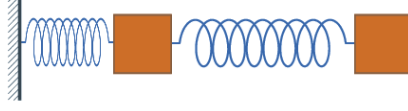## 4.1 2-spring systems with noise and friction



Figure 1: 2-spring systems

Consider a "Wall – linear stiff Spring – Mass – nonlinear soft Spring – Mass" system with both masses under isotropic noise and friction (Figure 1). The governing equations write as:

$$\begin{cases} dx & = p_x dt \\ dy & = p_y dt \\ dp_x & = -(\omega^2 x + (x-y)^3)dt - cp_x dt + \sigma dW_t^1 \\ dp_y & = -(y-x)^3 dt - cp_y dt + \sigma dW_t^2 \end{cases}$$

Note (1) this is a Langevin system with $H(x, y, p_x, p_y) = \frac{1}{2}p_x^2 + \frac{1}{2}p_y^2 + \frac{1}{2}\omega^2 x^2 + \frac{1}{4}(y-x)^4$ (2) $y$ is a slow variable but $x$ is not purely fast (there is a slow component in it).
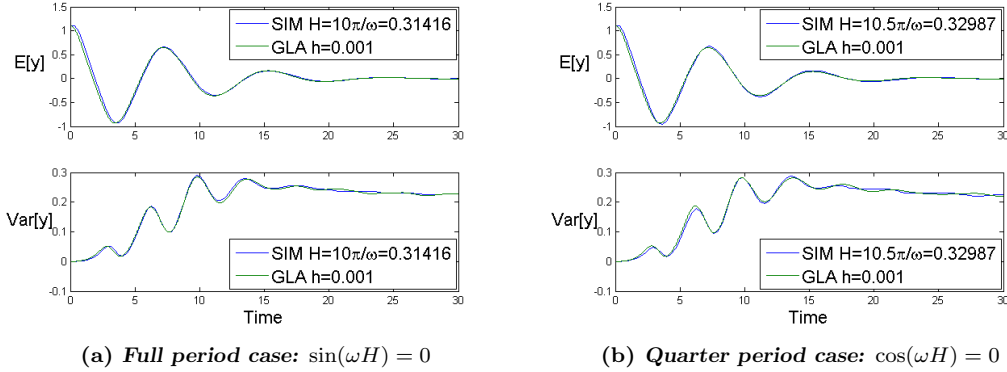


(a) **Full period case:** $\sin(\omega H) = 0$    (b) **Quarter period case:** $\cos(\omega H) = 0$

Figure 2: Empirical moments obtained by 1st-order SIM with large step $H$ and 1st-order GLA [3] with small step $h$. Parameters are $\omega = 100$, $c = 0.1$, $\beta = \frac{2c}{\sigma^2} = 10$, $x(0) = 0.8/\omega$, $y(0) = 1.1 + x(0)$, $p_x(0) = 0$, $p_y(0) = 0$; $h = 0.1/\omega$ and $H$ is chosen to be not scaling with $\omega$ yet corresponding to a resonant frequency; empirical moments are obtained by averaging 5000 simulations.

1st-order SIM (Integrator 2) is compared in Figure 2 to the benchmark of Geometric Langevin Integrator (GLA) [3] which is Boltzmann-Gibbs preserving and convergent. Agreements on empirical moments of integrated trajectories serve as evidences of structure preservation and convergence. The large timesteps used by SIM are chosen to be the resonance frequencies and they do produce stable accurate results. $O(\omega)$-fold acceleration is gained by SIM.
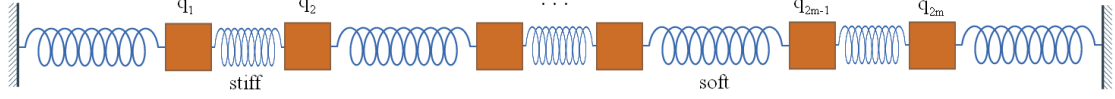
Figure 3: Fermi-Pasta-Ulam problem [11] – 1D chain of alternatively connected harmonic stiff and non-harmonic soft springs

## 4.2   Fermi-Pasta-Ulam problem

Consider the deterministic Fermi-Pasta-Ulam (FPU) problem [11] illustrated in Figure 3 and associated with the Hamiltonian

$$H(q,p) := \frac{1}{2}\sum_{i=1}^{m}(p_{2i-1}^2 + p_{2i}^2) + \frac{\omega^2}{4}\sum_{i=1}^{m}(q_{2i}-q_{2i-1})^2 + \sum_{i=0}^{m}(q_{2i+1}-q_{2i})^4 \tag{9}$$

Conventionally the following transformation is used

$$\begin{cases} x_i & = (q_{2i}+q_{2i-1})/\sqrt{2} \\ x_{m+i} & = (q_{2i}-q_{2i-1})/\sqrt{2} \\ y_i & = (p_{2i}+p_{2i-1})/\sqrt{2} \\ y_{m+i} & = (p_{2i}-p_{2i-1})/\sqrt{2} \end{cases} \quad i = 1,...m \tag{10}$$

so that the fast potential is diagonalized:

$$\begin{cases} H(x,y) & = \frac{1}{2}\sum_{i=1}^{2m} y_i^2 + V_f(x) + V_s(x) \\ V_f(x) & = \frac{\omega^2}{2}\sum_{i=1}^{m} x_{m+i}^2 \\ V_s(x) & = \frac{1}{4}((x_1-x_{m+1})^4 + \sum_{i=1}^{m-1}(x_{i+1}-x_{m+i+1}-x_i-x_{m+i})^4 + (x_m+x_{2m})^4) \end{cases}$$

The FPU problem is a well known benchmark problem [27, 16] for multiscale integrators because it exhibits different behaviors over widely separated timescales. The stiff springs (nearly) behave like harmonic oscillator with period $\sim O(\omega^{-1})$. Then the centres of masses linked by stiff springs (i.e. the middle points of stiff springs) change over a timescale $O(1)$. The third timescale $O(\omega)$ is associated with the rate of energy exchange among stiff springs. On the other hand, in addition to conservation of energy, the total energy of stiff springs behave almost like a constant. Comprehensive surveys on FPU problem, including discussions on timescales and numerical recipes, can be found in [16, 6].

We present in Figure 4 1st-order SIM simulation (Integrator 1) together with variational Euler (a.k.a. symplectic Euler) simulation of FPU over a time span of $O(\omega)$. Good results are obtained by SIM beyond the timescale of $O(1)$ (as guaranteed by Theorem 3.1) but actually over $O(\omega)$, and 200-fold ($\omega = 200$) acceleration is gained at the same time.

Notice that Mollified Impulse Methods with ShortAverage, LongAverage or LinearAverage filters [12] didn't accurately capture the rates of energy exchanging among stiff springs over $T = O(\omega)$ (results not shown).

## 5   Acknowledgement

(a) *1st-order SIM, large step* $H = 0.1$     (b) *Variational Euler, small step* $h = 0.1/\omega = 0.0005$
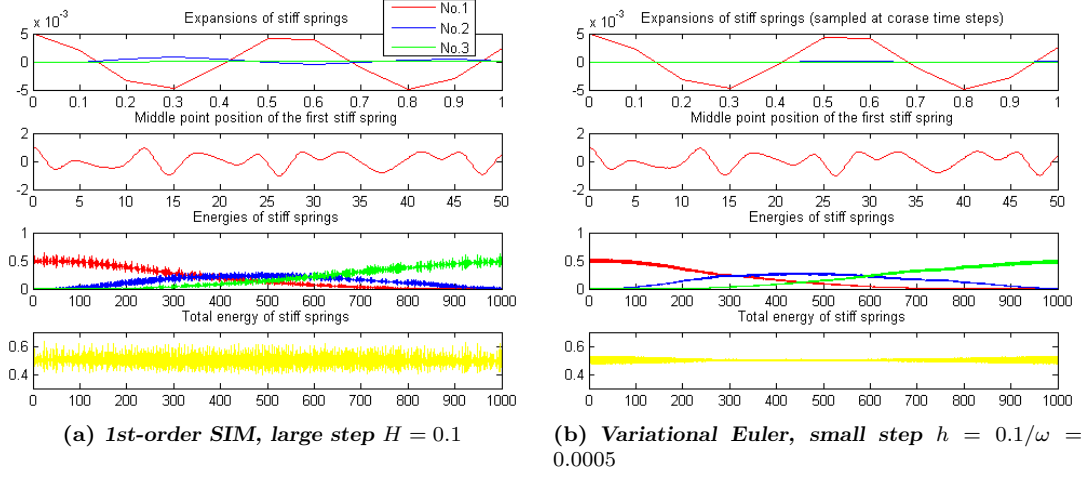
Figure 4: Simulations of FPU over $T = 5\omega$. Parameters are $\omega = 200$, $m = 3$, $x(0) = [1, 0, 0, 1/\omega, 0, 0]$, $y(0) = [0, 0, 0, 0, 0, 0]$. Different subplots use different time axes to accentuate different timescales: Subplot1 shows scaled expansions of three stiff springs $x_{m+i}$, which are fast variables; Subplot2 shows scaled middle point position of the first stiff spring $x_1$, which is one of the slow variables; Subplot3 shows the energy transferring pattern among stiff springs, which is even slower; Subplot4 shows the near-constant total energy of three stiff springs. The fast variables of stiff spring expansions are in fact oscillating much faster than shown in Subplots 1, for Subplots 1 are plotted by interpolating mesh points with a coarse mesh size of $H$.

# 6    Appendix

## 6.1    Proof of Theorem 3.1

Throughout this subsection Condition 3.1 is assumed. For a concise writing we also abuse the notation $O(x^n)$, which indicates some entity whose norm $\leq Cx^n$, where $C$ is a constant that doesn't change with $\epsilon$, i.e. doesn't depend on $\epsilon^{-1}$.

**Definition 6.1. Scaled energy norm:**

$$\Omega \triangleq \epsilon^{-1/2}\sqrt{K}$$
$$\| \begin{bmatrix} q \\ p \end{bmatrix} \|_E \triangleq \| \begin{bmatrix} q \\ \Omega^{-1}p \end{bmatrix} \|_2 = \sqrt{q^T q + \epsilon p^T K^{-1} p}$$

This is well defined because K is positive definite.

Since $\epsilon$ is very small, the following inequalities for converting between scaled energy norm and two-norm can be easily obtained:

**Proposition 6.1.** *Let* $x = \begin{bmatrix} q \\ p \end{bmatrix}$ *be any vector, then*

$$\epsilon^{1/2}\|\sqrt{K}\|_2^{-1}\|x\|_2 = \|\Omega\|_2^{-1}\|x\|_2 \leq \|x\|_E \leq \|x\|_2 \tag{11}$$

$$\| \begin{bmatrix} 0 \\ p \end{bmatrix} \|_E \leq \|\Omega^{-1}\|_2 \| \begin{bmatrix} 0 \\ p \end{bmatrix} \|_2 = \epsilon^{1/2}\|\sqrt{K}^{-1}\|_2\|x\|_2 \tag{12}$$

*Also, vector-norm-induced matrix norms satisfy*

$$\| \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \|_E \triangleq \sup \frac{\|Mx\|_E}{\|x\|_E} = \| \begin{bmatrix} M_{11} & M_{12}\Omega \\ \Omega^{-1}M_{21} & \Omega^{-1}M_{22}\Omega \end{bmatrix} \|_2 \tag{13}$$

9

**Lemma 6.1.** *Let $B(s) = \begin{bmatrix} B_{11}(s) & B_{12}(s) \\ B_{21}(s) & B_{22}(s) \end{bmatrix} = \exp(s \begin{bmatrix} 0 & I \\ -\epsilon^{-1}K & c \end{bmatrix})$, and $Rq(s)$ be the $Rq_k(H)$ defined in Integrator 2 with $H = s$ and arbitrary $k$, then*

$$\|B_{11}(s)\|_2 \leq 1 \tag{14}$$

$$\|B_{22}(s)\|_2 \leq 1 \tag{15}$$

$$\|B_{12}(s)\|_2 \leq |s| \tag{16}$$

$$\epsilon\|B_{21}(s)\|_2 \leq C_K|s| \tag{17}$$

$$\epsilon^{1/2}\|B_{11}(s) - I\|_2 \leq C_K|s| \tag{18}$$

$$\epsilon^{1/2}\|B_{22}(s) - I\|_2 \leq C_c|s| \tag{19}$$

$$\mathbb{E}\|Rq(s)\|_2^2 \leq \frac{1}{3}\|\sigma\|_2^2|s|^3 \tag{20}$$

$$\epsilon^{1/2}\|B(s) - I\|_2 \leq C_{Kc}|s| \tag{21}$$

$$\epsilon^{1/2}\|B(s) - I\|_E \leq C_{Kc}|s| \tag{22}$$

*where $C_K$, $C_c$ and $C_{Kc}$ are some positive real constants (may indicate different values in different inequalities), respectively dependent on $K$, $\sqrt{\epsilon}c$, $K$ and $\sqrt{\epsilon}c$ but independent of $\epsilon^{-1}$.*

*Proof.* Since $c$ and $K$ commute, they can be diagonalized simultaneously [18]. By the theory of linear ordinary differential equations [31], one can hence diagonalize $B_{11}$, $B_{12}$, $B_{21}$, $B_{22}$ simultaneously. Since each diagonal element can be investigated individually, assume without loss of generality that $\Omega = [\omega]_{ij} = \epsilon^{-1/2}\sqrt{K}$ and $c$ are both scalars, and use the notation of scalar $\omega$ and scalar $c$ thereafter.

Denote the damping ratio by $\zeta = \frac{c}{\omega}$. The solution to damped harmonic oscillator can be analytically obtained, and hence components of the flow operator $B_{11}, B_{12}, B_{21}, B_{22}$ as well.

When $\zeta < 1$ i.e. underdamping, which is usually the case since $\omega$ is large

$$B_{11}(s) = e^{-\omega\zeta s}(cos(\omega\sqrt{1-\zeta^2}s) + \frac{\zeta}{\sqrt{1-\zeta^2}}sin(\sqrt{1-\zeta^2}s)) \tag{23}$$

$$B_{12}(s) = \frac{e^{-\omega\zeta s}sin(\omega\sqrt{1-\zeta^2}s)}{\omega\sqrt{1-\zeta^2}} \tag{24}$$

$$B_{21}(s) = -\omega\frac{e^{-\omega\zeta s}sin(\omega\sqrt{1-\zeta^2}s)}{\sqrt{1-\zeta^2}} \tag{25}$$

$$B_{22}(s) = e^{-\omega\zeta s}(cos(\omega\sqrt{1-\zeta^2}s) - \frac{\zeta}{\sqrt{1-\zeta^2}}sin(\sqrt{1-\zeta^2}s)) \tag{26}$$

When $\zeta = 1$ i.e. critical damping,

$$B_{11}(s) = e^{-\omega s}(1 + \omega s) \tag{27}$$

$$B_{12}(s) = e^{-\omega s}s \tag{28}$$

$$B_{21}(s) = -\omega^2 e^{-\omega t}t \tag{29}$$

$$B_{22}(s) = e^{-\omega s}(1 - \omega s) \tag{30}$$

When $\zeta > 1$ i.e. over damping,

$$A(s) \triangleq e^{\omega s(-\zeta - \sqrt{\zeta^2 - 1})} \tag{31}$$

$$B(s) \triangleq e^{\omega s(-\zeta + \sqrt{\zeta^2 - 1})} \tag{32}$$

$$B_{11}(s) = \frac{\zeta(B - A) + \sqrt{\zeta^2 - 1}(A + B)}{2\sqrt{\zeta^2 - 1}} \tag{33}$$

$$B_{12}(s) = \frac{-A + B}{2\omega\sqrt{\zeta^2 - 1}} \tag{34}$$

$$B_{21}(s) = \frac{\omega(A - B)}{2\sqrt{\zeta^2 - 1}} \tag{35}$$

$$B_{22}(s) = \frac{\zeta(A - B) + \sqrt{\zeta^2 - 1}(A + B)}{2\sqrt{\zeta^2 - 1}} \tag{36}$$

$$\tag{37}$$

By routine investigations on local extremes using calculus, it can be shown in all three cases that

$$\|B_{11}(s)\|_2 \leq 1 \tag{38}$$

$$\|B_{22}(s)\|_2 \leq 1 \tag{39}$$

$$\|B_{12}(s)\|_2 \leq s \tag{40}$$

$$\|B_{21}(s)\|_2 \leq \omega^2 s \tag{41}$$

$$\|B_{11}(s) - I\|_2 \leq \omega s \tag{42}$$

$$\|B_{22}(s) - I\|_2 \leq \begin{cases} \omega s & \zeta \leq 1 \\ 2\zeta\omega s & \zeta > 1 \end{cases} \tag{43}$$

When $\zeta > 1$, since $c = O(\epsilon^{-1/2})$ (Condition 3.1), $2\zeta\omega s = O(\epsilon^{-1/2})s$. Therefore $\epsilon^{1/2}\|B_{22}(s) - I\|_2 \leq C_c|s|$ always holds.

Also,

$$\mathbb{E}\|Rq(s)\|_2^2 = \mathbb{E}\|\int_0^s B_{12}(t)\sigma dW_t\|_2^2$$

$$= \int_0^s \|\sigma B_{12}(t)\|_2^2 dt \leq \frac{1}{3}\|\sigma\|_2^2|s|^3 \tag{44}$$

For a proof on norm bounds of the entire matrice we use only bounds of dimensionless block elements:

$$\|B - I\|_2 = \|\begin{bmatrix} B_{11} - I & B_{12} \\ B_{21} & B_{22} - I \end{bmatrix}\|_2 \tag{45}$$

$$\leq \|\begin{bmatrix} \Omega & 0 \\ 0 & \Omega \end{bmatrix}\|_2 \|\begin{bmatrix} \Omega^{-1}(B_{11} - I) & \Omega^{-1}B_{12} \\ \Omega^{-1}B_{21} & \Omega^{-1}(B_{22} - I) \end{bmatrix}\|_2 \tag{46}$$

$$= \epsilon^{-1/2}\|\begin{bmatrix} O(s) & \epsilon^{1/2}O(s) \\ \epsilon^{-1/2}O(s) & O(s) \end{bmatrix}\|_2 \tag{47}$$

It's easy to prove that for any scalar $a$

$$\|\begin{bmatrix} M_{11} & aM_{12} \\ M_{21} & M_{22} \end{bmatrix}\|_2 = \|\begin{bmatrix} M_{11} & M_{12} \\ aM_{21} & M_{22} \end{bmatrix}\|_2 \tag{48}$$

11

Therefore

$$\|B - I\|_2 = \epsilon^{-1/2}\| \begin{bmatrix} O(s) & O(s) \\ O(s) & O(s) \end{bmatrix} \|_2 = \epsilon^{-1/2}O(s) \tag{49}$$

Similarly,

$$
\begin{aligned}
\|B - I\|_E &= \epsilon^{-1/2}\| \begin{bmatrix} B_{11} - I & B_{12}\Omega \\ \Omega^{-1}B_{21} & \Omega^{-1}B_{22}\Omega - I \end{bmatrix} \|_2 \\
&\leq \epsilon^{-1/2}\| \begin{bmatrix} O(s) & \epsilon^{1/2}\epsilon^{-1/2}O(s) \\ \epsilon^{-1/2}\epsilon^{1/2}O(s) & \epsilon^{1/2}O(s)\epsilon^{-1/2} \end{bmatrix} \|_2 \\
&= \epsilon^{-1/2}\| \begin{bmatrix} O(s) & O(s) \\ O(s) & O(s) \end{bmatrix} \|_2 \\
&= \epsilon^{-1/2}O(s) \tag{50}
\end{aligned}
$$

□

**Remark 6.1.** In the special case of $c = 0$, bounds of block elements can be easily obtained since

$$|cos(\omega s)| \leq 1$$

$$\epsilon K^{-1}|-\omega sin(\omega s)| = |\frac{-\omega sin(\omega s)}{\omega^2}| \leq |s|$$

$$\epsilon^{1/2}\sqrt{K}^{-1}|cos(\omega s) - 1| = |-2sin^2(\omega s/2)/\omega| \leq |-2sin(\omega s/2)/\omega| \leq |s|$$

**Lemma 6.2.** *The solution to SDE $dX = AXdt + f(X)dt + \Sigma dW_t$ can be written in the following integral form:*

$$X(t) = e^{At}X(0) + \int_0^t e^{A(t-s)}f(X(s))ds + \int_0^t e^{A(t-s)}\Sigma dW_s \tag{51}$$

*Proof.* Let $Y(t) = e^{-At}X(t)$, then by Ito's formula and $dX = AXdt + f(X)dt + \Sigma dW_t$

$$dY = e^{-At}f(X(t))dt + e^{-At}\Sigma dW_t \tag{52}$$

This in the integral form is

$$Y(t) = Y(0) + \int_0^t e^{-As}f(X(s))ds + \int_0^t e^{-As}\Sigma dW_s \tag{53}$$

Hence

$$X(t) = e^{At}X(0) + \int_0^t e^{A(t-s)}f(X(s))ds + \int_0^t e^{A(t-s)}\Sigma dW_s \tag{54}$$

□

**Lemma 6.3.** *Consider two continuous stochastic dynamical systems, the original dynamics and the bridge dynamics:*

$$\begin{cases} dq &= pdt \\ dp &= -\epsilon^{-1}Kqdt - \nabla V(q)dt - cpdt + \sigma dW_t \\ q(0) &= q_0 \\ p(0) &= p_0 \end{cases} \tag{55}$$

$$\begin{cases} d\tilde{q} &=& \tilde{p}dt \\ d\tilde{p} &=& -\epsilon^{-1}K\tilde{q}dt - \nabla V(q_0)dt - c\tilde{p}dt + \sigma dW_t \\ \tilde{q}(0) &=& q_0 \\ \tilde{p}(0) &=& p_0 \end{cases} \tag{56}$$

Then $(\mathbb{E}\| \begin{bmatrix} \tilde{q}(H) \\ \tilde{p}(H) \end{bmatrix} - \begin{bmatrix} q(h) \\ p(h) \end{bmatrix} \|_E^2)^{1/2} \le C|H|^{3/2}$, where $C$ is a positive constant independent of $\epsilon^{-1}$ but dependent on the scaleless elasticity matrix $K$, scaled damping coefficient $\sqrt{\epsilon}c$, amplitude of noise $\sigma$, and slow potential $V(\cdot)$.

*Proof.* Rewrite the original dynamics (55) as

$$\begin{cases} dq &=& pdt \\ dp &=& \epsilon^{-1}Kqdt - \nabla V(q_0)dt + (\nabla V(q_0) - \nabla V(q))dt - cpdt + \sigma dW_t \\ q(0) &=& q_0 \\ p(0) &=& p_0 \end{cases} \tag{57}$$

Let $x(t) = \begin{bmatrix} q(t) \\ p(t) \end{bmatrix}$, $\tilde{x}(t) = \begin{bmatrix} \tilde{q}(t) \\ \tilde{p}(t) \end{bmatrix}$, $B(t) = \exp(t\begin{bmatrix} 0 & I \\ -\epsilon^{-1}K & -c \end{bmatrix})$, $b = \begin{bmatrix} 0 \\ -\nabla V(q_0) \end{bmatrix}$, $g(q,p) = g(x) = \begin{bmatrix} 0 \\ \nabla V(q_0) - \nabla V(q) \end{bmatrix}$, and $\Sigma = \begin{bmatrix} 0 \\ \sigma \end{bmatrix}$. Then by Lemma 6.2 solutions to the original dynamics and bridge dynamics can be respectively written as:

$$\begin{aligned} x(t) &=& B(t)x(0) + \int_0^t B(t-s)bds + \int_0^t B(t-s)\Sigma dW_s + \int_0^t B(t-s)g(x(s))ds \\ \tilde{x}(t) &=& B(t)x(0) + \int_0^t B(t-s)bds + \int_0^t B(t-s)\Sigma dW_s \end{aligned} \tag{58}$$

Notice for any vector $y$ and positive $t$ that $\|B(t)y\|_E \le \|y\|_E$, because energy is decaying in the system $\ddot{q} + c\dot{q} + \epsilon^{-1}Kq = 0$. Together with Cauchy-Schwarz we have

$$\begin{aligned} \mathbb{E}\|\tilde{x}(t) - x(t)\|_E^2 &=& \mathbb{E}\|\int_0^t B(t-s)g(x(s))ds\|_E^2 \\ &\le& t\int_0^t \mathbb{E}\|B(t-s)g(x(s))\|_E^2 ds \\ &\le& t\int_0^t \mathbb{E}\|g(x(s))\|_E^2 ds \end{aligned} \tag{59}$$

By Condition 3.1, assume $\nabla V(\cdot)$ is Lipschitz continuous with coefficient $L$, then almost surely

$$\begin{aligned} \|g(x(s))\|_E &=& \| \begin{bmatrix} 0 \\ \nabla V(q_0) - \nabla V(q(s)) \end{bmatrix} \|_E \\ &\le& \sqrt{\epsilon}\|\sqrt{K}^{-1}\|_2\|\nabla V(q_0) - \nabla V(q(s))\|_2 \\ &\le& L\sqrt{\epsilon}\|\sqrt{K}^{-1}\|_2\|q(s) - q_0\|_2 \end{aligned} \tag{60}$$

Similarly, since

$$x(t) = B(t)x(0) + \int_0^t B(t-s)\begin{bmatrix} 0 \\ -\nabla V(q(s)) \end{bmatrix}ds + \int_0^t B(t-s)\Sigma dW_s \tag{61}$$

13

we have

$$\|x(s) - B(s)x_0 - \int_0^s B(s-t)\Sigma dW_t\|_2 \leq \int_0^s \|\nabla V(q(t))\|_2 dt \tag{62}$$

By Condition 3.1, $\nabla V(\cdot)$ is bounded, and hence the above is $O(s)$.

We now can bound (60) and therefore (59) with the aid of (62) and Lemma 6.1:

$$
\begin{aligned}
&\mathbb{E}\|q(s) - q_0\|_2^2 \\
\leq\ & \mathbb{E}\|x(s) - x_0\|_2^2 \\
\leq\ & \mathbb{E}\left(\|x(s) - B(s)x_0 - \int_0^s B(t-s)\Sigma dW_s\|_2 + \|B(s)x_0 - x_0\|_2 + \|\int_0^s B(t-s)\Sigma dW_s\|_2\right)^2 \\
\leq\ & 3\mathbb{E}\left(\|x(s) - B(s)x_0 - \int_0^s B(t-s)\Sigma dW_s\|_2^2 + \|B(s)x_0 - x_0\|_2^2 + \|\int_0^s B(t-s)\Sigma dW_s\|_2^2\right) \\
=\ & 3\left(O(s^2) + \epsilon^{-1}O(s^2)\mathbb{E}\|x_0\|_2^2 + \int_0^s \sigma^2(B_{12}(t-s)^2 + B_{22}(t-s)^2)dt\right) \\
=\ & O(s^2) + \epsilon^{-1}O(s^2)\mathbb{E}\|x_0\|_2^2 + O(s^3) + O(s) \tag{63}
\end{aligned}
$$

By Condition 3.1, $\mathbb{E}\|x_0\|_2^2 = O(1)$. Therefore, the above expression is $\epsilon^{-1}O(s^2) + O(s)$.

This gives $\mathbb{E}\|g(\tilde{x}(s))\|_E^2 = O(s)$ independent of $\epsilon^{-1}$, and eventually $\mathbb{E}\|\tilde{x}(h) - x(h)\|_E^2 = O(h^3)$. $\qquad\square$

**Lemma 6.4.** *Consider the discrete stochastic dynamical system given by 1st-order SIM (Integrator 2):*

$$
\begin{cases}
q_H = B_{11}(H)q_0 + B_{12}(H)p_0 + Rq(H) \\
p_H = B_{21}(H)q_0 + B_{22}(H)p_0 + Rp(H) - H\nabla V(B_{11}(H)q_0 + B_{12}(H)p_0 + Rq(H))
\end{cases} \tag{64}
$$

*Then a comparison with bridge dynamics (56) gives $\mathbb{E}\|q_H - \tilde{q}(H)\|_2^2 \leq CH^4$ and $\mathbb{E}\|\Omega^{-1}(p_H - \tilde{p}(H))\|_2^2 \leq CH^4$, and therefore*

$$\left(\mathbb{E}\|\begin{bmatrix} q_H \\ p_H \end{bmatrix} - \begin{bmatrix} \tilde{q}(H) \\ \tilde{p}(H) \end{bmatrix}\|_E^2\right)^{1/2} \leq CH^2 \tag{65}$$

*where $C$'s are positive constants independent of $\epsilon^{-1}$ but dependent on scaleless elasticity matrix $K$, scaled damping coefficient $\sqrt{\epsilon}c$, amplitude of noise $\sigma$, and slow potential $V(\cdot)$.*

*Proof.* The exact solution to the bridge dynamics is

$$
\begin{cases}
\tilde{q}(H) = B_{11}(H)q_0 + B_{12}(H)p_0 + \int_0^H B_{12}(s)(-\nabla V(q_0))ds + Rq(H) \\
\tilde{p}(H) = B_{21}(H)q_0 + B_{22}(H)p_0 + \int_0^H B_{22}(s)(-\nabla V(q_0))ds + Rp(H)
\end{cases} \tag{66}
$$

Hence almost surely $\tilde{q}(H) - q_H = \int_0^H B_{12}(s)(-\nabla V(q_0))ds$.

Since $B_{12}(s) = O(s)$ by Lemma 6.1, and $\mathbb{E}\| - \nabla V(q_0)\|_2^2$ is bounded by Condition (3.1), one gets

$$
\begin{aligned}
\mathbb{E}\|\tilde{q}(H) - q_H\|_2^2 &\leq H\int_0^H \mathbb{E}\|B_{12}(s)(-\nabla V(q_0))\|_2^2 ds \\
&\leq H\int_0^H \mathbb{E}(\|B_{12}(s)\|_2\| - \nabla V(q_0)\|_2)^2 ds \\
&= H\int_0^H O(s^2)\mathbb{E}\| - \nabla V(q_0)\|_2^2 ds \\
&= O(H^4) \tag{67}
\end{aligned}
$$

14

Investigation on $p$ by applying Lemma 6.1 and Condition 3.1 gives:

$$\mathbb{E}\|\Omega^{-1}(\tilde{p}(H) - p_H)\|_2^2$$

$$= \mathbb{E}\|\Omega^{-1}(\int_0^H B_{22}(s)ds(-\nabla V(q_0)) + H\nabla V(B_{11}(H)q_0 + B_{12}(H)p_0 + Rq(H)))\|_2^2$$

$$= \mathbb{E}\|\int_0^H \Omega^{-1}(B_{22}(s) - I)ds(-\nabla V(q_0)) + H\Omega^{-1}(\nabla V(B_{11}(H)q_0 + B_{12}(H)p_0$$
$$+ Rq(H)) - \nabla V(q_0))\|_2^2$$

$$\leq 2\mathbb{E}[\|\int_0^H \Omega^{-1}(B_{22}(s) - I)ds(-\nabla V(q_0))\|_2^2 + \|H\Omega^{-1}(\nabla V(B_{11}(H)q_0 + B_{12}(H)p_0$$
$$+ Rq(H)) - \nabla V(q_0))\|_2^2]$$

$$\leq 2[H\int_0^H \mathbb{E}\|\Omega^{-1}(B_{22}(s) - I)(-\nabla V(q_0))\|_2^2 ds + \mathbb{E}\|H\Omega^{-1}(\nabla V(B_{11}(H)q_0$$
$$+ B_{12}(H)p_0 + Rq(H)) - \nabla V(q_0))\|_2^2]$$

$$\leq 2H[\int_0^H \|\Omega^{-1}(B_{22}(s) - I)\|_2^2 ds\mathbb{E}\|(-\nabla V(q_0))\|_2^2 + H\mathbb{E}\|\Omega^{-1}(\nabla V(B_{11}(H)q_0$$
$$+ B_{12}(H)p_0 + Rq(H)) - \nabla V(q_0))\|_2]$$

$$\leq 2H[O(H^3) + L^2 H\mathbb{E}\|\Omega^{-1}(B_{11}(H)q_0 + B_{12}(H)p_0 + Rq(H) - q_0)\|_2^2]$$

$$\leq 2H[O(H^3) + 3L^2 H(\mathbb{E}\|\Omega^{-1}(B_{11}(H) - I)q_0\|_2^2 + \|\Omega^{-1}B_{12}(H)p_0\|_2^2$$
$$+ \mathbb{E}\|\Omega^{-1}Rq(H)\|_2^2)]$$

$$\leq 2H[O(H^3) + 3L^2 H(\|\Omega^{-1}(B_{11}(H) - I)\|_2^2 \mathbb{E}\|q_0\|_2^2 + \|B_{12}(H)\|_2^2 \mathbb{E}\|p_0\|_2^2$$
$$+ \mathbb{E}\|Rq(H)\|_2^2)]$$

$$\leq 2H[O(H^3) + 3L^2 H(O(H)^2\mathbb{E}\|q_0\|_2^2 + O(H)^2\mathbb{E}\|p_0\|_2^2 + O(H^3))]$$

$$= O(H^4) \tag{68}$$

Therefore $\mathbb{E}\|\begin{bmatrix} q_H \\ p_H \end{bmatrix} - \begin{bmatrix} \tilde{q}(H) \\ \tilde{p}(H) \end{bmatrix}\|_E^2 = O(H^4)$ independent of $\epsilon^{-1}$. $\qquad\square$

**Lemma 6.5.** *Consider evolutions of different local initial conditions under the bridge dynamics:*

$$\begin{cases} d\tilde{q}_1 &= \tilde{p}_1 dt \\ d\tilde{p}_1 &= -\epsilon^{-1}K\tilde{q}_1 dt - \nabla V(\tilde{q}_1(0))dt - c\tilde{p}_1 dt + \sigma dW_t \end{cases} \tag{69}$$

$$\begin{cases} d\tilde{q}_2 &= \tilde{p}_2 dt \\ d\tilde{p}_2 &= -\epsilon^{-1}K\tilde{q}_2 dt - \nabla V(\tilde{q}_2(0))dt - c\tilde{p}_2 dt + \sigma dW_t \end{cases} \tag{70}$$

*Denote by $L$ the Lipschitz coefficient of $\nabla V(\cdot)$ (i.e. $\|\nabla V(a) - \nabla V(b)\|_2 \leq L\|a - b\|_2$), then almost surely*

$$\|\begin{bmatrix} \tilde{q}_1(H) - \tilde{q}_2(H) \\ \tilde{p}_1(H) - \tilde{p}_2(H) \end{bmatrix}\|_E \leq (1 + HL)\|\begin{bmatrix} \tilde{q}_1(0) - \tilde{q}_2(0) \\ \tilde{p}_1(0) - \tilde{p}_2(0) \end{bmatrix}\|_E \tag{71}$$

*Proof.* Write out the solution to the bridge dynamics in integral form:

$$
\begin{bmatrix} \tilde{q}_1(H) \\ \tilde{p}_1(H) \end{bmatrix} = B(H) \begin{bmatrix} \tilde{q}_1(0) \\ \tilde{p}_1(0) \end{bmatrix} + \int_0^H B(H-s) \begin{bmatrix} 0 \\ -\nabla V(\tilde{q}_1(0)) \end{bmatrix} ds + \int_0^H B(H-s)\Sigma dW_s
$$

$$
\begin{bmatrix} \tilde{q}_2(H) \\ \tilde{p}_2(H) \end{bmatrix} = B(H) \begin{bmatrix} \tilde{q}_2(0) \\ \tilde{p}_2(0) \end{bmatrix} + \int_0^H B(H-s) \begin{bmatrix} 0 \\ -\nabla V(\tilde{q}_2(0)) \end{bmatrix} ds + \int_0^H B(H-s)\Sigma dW_s
$$

(72)

Hence almost surely

$$
\| \begin{bmatrix} \tilde{q}_1(H) - \tilde{q}_2(H) \\ \tilde{p}_1(H) - \tilde{p}_2(H) \end{bmatrix} \|_E
$$

$$
\leq \quad \|B(H) \begin{bmatrix} \tilde{q}_1(0) - \tilde{q}_2(0) \\ \tilde{p}_1(0) - \tilde{p}_2(0) \end{bmatrix} \|_E + \int_0^H \|B(H-s) \begin{bmatrix} 0 \\ \nabla V(\tilde{q}_2(0)) - \nabla V(\tilde{q}_1(0)) \end{bmatrix} \|_E ds
$$

$$
\leq \quad \| \begin{bmatrix} \tilde{q}_1(0) - \tilde{q}_2(0) \\ \tilde{p}_1(0) - \tilde{p}_2(0) \end{bmatrix} \|_E + H \| \begin{bmatrix} 0 \\ \nabla V(\tilde{q}_2(0)) - \nabla V(\tilde{q}_1(0)) \end{bmatrix} \|_E
$$

$$
\leq \quad \| \begin{bmatrix} \tilde{q}_1(0) - \tilde{q}_2(0) \\ \tilde{p}_1(0) - \tilde{p}_2(0) \end{bmatrix} \|_E + HL \| \begin{bmatrix} 0 \\ \tilde{q}_2(0) - \tilde{q}_1(0) \end{bmatrix} \|_E
$$

$$
\leq \quad \| \begin{bmatrix} \tilde{q}_1(0) - \tilde{q}_2(0) \\ \tilde{p}_1(0) - \tilde{p}_2(0) \end{bmatrix} \|_E + HL \| \begin{bmatrix} \tilde{q}_2(0) - \tilde{q}_1(0) \\ 0 \end{bmatrix} \|_E
$$

$$
\leq \quad (1 + HL) \| \begin{bmatrix} \tilde{q}_1(0) - \tilde{q}_2(0) \\ \tilde{p}_1(0) - \tilde{p}_2(0) \end{bmatrix} \|_E
$$

(73)

$\square$

**Remark 6.2.** If the traditional method of investigating the Lipschitz coefficient of the vector field is employed to evolve the separation of local initial conditions, $\epsilon^{-1}$ will exhibit in the bound of separation. Instead we only looked at the soft part of the vector field and whence obtained a uniform bound.

**Theorem 3.1 (global error bound in energy norm).**

*Proof.*
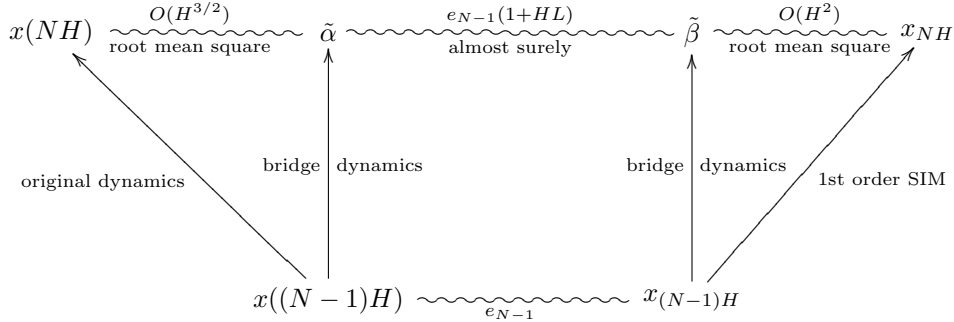


**Atlas of error propagation**

Let $e_N = (\mathbb{E}\|x(NH) - x_{NH}\|_E^2)^{1/2}$. Let $\tilde{\alpha}$ and $\tilde{\beta}$ be respectively the evolution of the real solution $x((N-1)H)$ and the numerical solution $x_{(N-1)H}$ by time $H$ under the bridge dynamics (56).

16

Then by Lemma 6.3 and 6.4, there exist constants $C_1$ and $C_2$ independent of $\epsilon^{-1}$ such that

$$
\begin{aligned}
(\mathbb{E}\|X(NH) - \tilde{\alpha}\|_E^2)^{1/2} &\leq C_1 H^{3/2} \\
(\mathbb{E}\|\tilde{\beta} - X_{Nh}\|_E^2)^{1/2} &\leq C_2 H^2
\end{aligned}
\tag{74}
$$

Also since $\|\tilde{\alpha} - \tilde{\beta}\|_E \leq (1 + HL)\|x((N-1)h) - x_{(N-1)h}\|_E$ almost surely (Lemma 6.5), we have:

$$
(\mathbb{E}\|\tilde{\alpha} - \tilde{\beta}\|_E^2)^{1/2} \leq (1 + HL)e_{N-1}
\tag{75}
$$

All in all,

$$
\begin{aligned}
e_N &\leq (\mathbb{E}\|x(NH) - \tilde{\alpha}\|_E^2)^{1/2} + (\mathbb{E}\|\tilde{\alpha} - \tilde{\beta}\|_E^2)^{1/2} + (\mathbb{E}\|\tilde{\beta} - X_{NH})\|_E^2)^{1/2} \\
&\leq (1 + HL)e_{N-1} + (C_1 + C_2)H^{3/2} \\
&= (1 + HL)^N e_0 + (C_1 + C_2)H^{3/2}\frac{(1 + HL)^N - 1}{(1 + HL) - 1} \\
&\leq (C_1 + C_2)H^{1/2}\frac{e^{NHL} - 1}{L} = \frac{(C_1 + C_2)(e^{TL} - 1)}{L}H^{1/2}
\end{aligned}
\tag{76}
$$

Therefore letting $C = \frac{(C_1+C_2)(e^{TL}-1)}{L}$ we have

$$
(\mathbb{E}\|q(T) - q_T\|_2^2)^{1/2} \leq e_N \leq CH^{1/2}
\tag{77}
$$

$$
(\mathbb{E}\|p(T) - p_T\|_2^2)^{1/2} \leq \epsilon^{-1/2}\|\sqrt{K}\|_2 e_N \leq \epsilon^{-1/2}\|\sqrt{K}\|_2 CH^{1/2}
\tag{78}
$$

$\square$

# References

[1] G. ARIEL, B. ENGQUIST, AND Y.-H. TSAI, *A multiscale method for highly oscillatory ordinary differential equations with resonance*, Math. Comput., 78 (2009), p. 929.

[2] ——, *A reversible multiscale integration method*, To appear, Comm. Math. Sci., (2009).

[3] N. BOU-RABEE AND H. OWHADI, *Long-run accuracy of variational integrators in the stochastic context*, SIAM J. Numer. Anal., 48 (2010), pp. 278–297.

[4] M. CALVO AND J. SANZ-SERNA, *Heterogeneous multiscale methods for mechanical systems with vibrations.*, preprint, (2008).

[5] ——, *Instabilities and inaccuracies in the integration of highly oscillatory problems*, SIAM J. Sci. Comput., 31 (2009), pp. 1653–1677.

[6] D. COHEN, T. JAHNKE, K. LORENZ, AND C. LUBICH, *Numerical integrators for highly oscillatory Hamiltonian systems: a review*, in Analysis, modeling and simulation of multiscale problems, Springer, Berlin, 2006, pp. 553–576.

[7] W. E, *Analysis of the heterogeneous multiscale method for ordinary differential equations*, Commun. Math. Sci., 1 (2003), pp. 423–436.

[8] W. E, B. ENGQUIST, X. LI, W. REN, AND E. VANDEN-EIJNDEN, *Heterogeneous multiscale methods: a review*, Commun. Comput. Phys., 2 (2007), pp. 367–450.

[9] W. E, W. Ren, and E. Vanden-Eijnden, *A general strategy for designing seamless multiscale methods*, J. Comput. Phys., 228 (2009), pp. 5437–5453.

[10] B. Engquist and Y.-H. Tsai, *Heterogeneous multiscale methods for stiff ordinary differential equations*, Math. Comp., 74 (2005), pp. 1707–1742 (electronic).

[11] E. Fermi, J. Pasta, and S. Ulam, *Studies of nonlinear problems*, Tech. Report LA-1940, Los Alamos Scientific Laboratory, 1955.

[12] B. García-Archilla, J. Sanz-Serna, and R. Skeel, *Long-time-step methods for oscillatory differential equations*, SIAM J. Sci. Comput., 20 (3) (1999), pp. 930–963.

[13] D. Givonand, I. G. Kevrekidis, and R. Kupferman, *Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems*, Commun. Math. Sci., 4 (2006), pp. 707–729.

[14] V. Grimm and M. Hochbruck, *Error analysis of exponential integrators for oscillatory second-order differential equations*, J. Phys. A: Math. Gen., 39 (2006), pp. 5495–5507.

[15] H. Grubmuller, H. Heller, A. Windemuth, and K. Schulten, *Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions*, Mol. Sim., 6 (1991), pp. 121–142.

[16] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer, Heidelberg Germany, second ed., 2004.

[17] M. Hochbruck and C. Lubich, *A gautschi-type method for oscillatory second-order differential equations*, Numer. Math., 83 (1999), pp. 403–426.

[18] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[19] I. Kevrekidis, C. Gear, J. Hyman, P. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis*, Commun. Math. Sci., 1 (2003), pp. 715–762.

[20] I. Kevrekidis and G. Samaey, *Equation-free multiscale computation: Algorithms and applications*, Annual Review of Physical Chemistry, 60 (2009), pp. 321–344. PMID: 19335220.

[21] P. Lax and R. Richtmyer, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math., 9 (1956), pp. 267–293.

[22] C. Le Bris and F. Legoll, *Integrators for highly oscillatory Hamiltonian systems: an homogenization approach*, Tech. Report 6252, Inria Rapport de recherche, 2007.

[23] B. Leimkuhler and S. Reich, *A reversible averaging integrator for multiple time-scale dynamics*, J. Comput. Phys., 171 (2001), pp. 95–114.

[24] T. Li, A. Abdulle, and W. E, *Effectiveness of implicit methods for stiff stochastic differential equations*, Commun. Comput. Phys., 3 (2008), pp. 295–307.

[25] G. I. Marchuk, *Some applications of splitting-up methods to the solution of mathematical physics problems*, Applik. Mat., 13 (1968), pp. 103–132.

[26] J. E. Marsden and M. West, *Discrete mechanics and variational integrators*, Acta Numerica, (2001), pp. 357–514.

[27] R. McLachlan and D. OíNeale, *Comparison of integrators for the Fermi-Pasta-Ulam problem.*, preprint NI07052-HOP, Isaac Newton Institute for Mathematical Sciences., (2007). http://www.newton.ac.uk/preprints/NI07052.pdf.

[28] R. McLachlan and M. Perlmutter, *Conformal Hamiltonian systems*, J. Geom. Phys., 39 (2001), pp. 276–300.

[29] G. Milstein and M. Tretyakov, *Quasi-symplectic methods for Langevin-type equations*, IMA J. Numer. Anal., 23 (2003), pp. 593–626.

[30] F. Neri, *Lie algebras and canonical integration*, tech. report, Department of Physics, University of Maryland, 1988.

[31] L. Perko, *Differential equations and dynamical systems*, Springer, 2001.

[32] J. Sanz-Serna, *Mollified impulse methods for highly oscillatory differential equations*, SIAM J. Numer. Anal., 46 (2) (2008), pp. 1040–1059.

[33] J. Sanz-Serna, G. Ariel, and Y.-H. Tsai, *Multiscale methods for stiff and constrained mechanical systems.*, preprint, (2009).

[34] R. D. Skeel and J. A. Izaguirre, *An impulse integrator for langevin dynamics*, Mol. Phys., 100 (2002), pp. 3885–3891.

[35] A. Stern and E. Grinspun, *Implicit-explicit variational integration of highly oscillatory problems*, Multiscale Model. Simul., 7 (2009), pp. 1779–1794.

[36] G. Strang, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal., 5 (1968), pp. 506–517.

[37] M. Tao, H. Owhadi, and J. E. Marsden, *Non-intrusive and structure preserving multiscale integration of stiff ODEs, SDEs and Hamiltonian systems with hidden slow dynamics via flow averaging*, (2010). arXiv:0908.1241v2. Accepted by SIAM MMS.

[38] H. F. Trotter, *Product of semigroups of operators*, Proc. Amer. Math. Soc., 10 (1959), pp. 545–551.

[39] M. Tuckerman, B. J. Berne, and G. J. Martyna, *Reversible multiple time scale molecular dynamics*, J. Chem. Phys., 97 (1992), pp. 1990–2001.

[40] G. Zhang and T. Schlick, *LIN: A new algorithm to simulate the dynamics of biomolecules by combining implicit-integration and normal mode techniques*, J. Comp. Chem., 14 (1993), pp. 1212–1233.